```
                                                         GGCACGAGTCGGAGCCGGG

CGGAGGGGAGGGGGGAAAGAGGAGCGCAGGGTGAGAGTGAGCCGCAGGCTTCGGGAGGCGAGGGGGCGGGGGGAGCAGC

                                                                M   P   G   E      4
GCCGAGGYCGCCGCCTCCGCCTCCGCCGCCTAGGACTAGGGGGTGGGGACGGACAAGCCCG ATG CCG GGG GAG     12

 T   E   E   P   R   P   P   E   Q   Q   D   Q   E   G   G   E   A   A   K   A      24
ACG GAA GAG CCG AGA CCC CCG GAG CAG CAG GAC CAG GAA GGG GGA GAG GCG GCC AAG GCG    72

 A   P   E   E   P   Q   Q   R   P   P   E   A   V   A   A   A   P   A   G   T      44
GCT CCG GAG GAG CCC CAA CAA CGG CCC CCT GAG GCG GTC GCG GCG GCG CCT GCA GGG ACC   132

 T   S   S   R   V   L   R   G   G   R   D   R   G   R   A   A   A   A   A   A      64
ACT AGC AGC CGC GTG CTG AGG GGA GGT CGG GAC CGA GGC CGG GCC GCT GCG GCC GCC GCC   192

 A   A   A   V   S   R   R   R   K   A   E   Y   P   R   R   R   R   S   S   P      84
GCC GCA GCT GTG TCC CGC CGG AGG AAG GCC GAG TAT CCC CGC CGG CGG AGG AGC AGC CCC   252

 S   A   R   P   P   D   V   P   G   Q   Q   P   Q   A   A   K   S   P   S   P     104
AGC GCC AGG CCT CCC GAC GTC CCC GGG CAG CAG CCC CAG GCC GCG AAG TCC CCG TCT CCA   312

 V   Q   G   K   K   S   P   R   L   L   C   I   E   K   V   T   T   D   K   D     124
GTT CAG GGC AAG AAG AGT CCG CGA CTC CTA TGC ATA GAA AAA GTA ACA ACT GAT AAA GAT   372

 P   K   E   E   K   E   E   E   D   D   S   A   L   P   Q   E   V   S   I   A     144
CCC AAG GAA GAA AAA GAG GAA GAA GAC GAT TCT GCC CTC CCT CAG GAA GTT TCC ATT GCT   432

 A   S   R   P   S   R   G   W   R   S   S   R   T   S   V   S   R   H   R   D     164
GCA TCT AGA CCT AGC CGG GGC TGG CGT AGT AGT AGG ACA TCT GTT TCT CGC CAT CGT GAT   492

 T   E   N   T   R   S   S   R   S   K   T   G   S   L   Q   L   I   C   K   S     184
ACA GAG AAC ACC CGA AGC TCT CGG TCC AAG ACC GGT TCA TTG CAG CTC ATT TGC AAG TCA   552

 E   P   N   T   D   Q   L   D   Y   D   V   G   E   E   H   Q   S   P   G   G     204
GAA CCA AAT ACA GAC CAA CTT GAT TAT GAT GTT GGA GAA GAG CAT CAG TCT CCA GGT GGC   612

 I   S   G   E   E   E   E   E   E   E   E   M   L   I   S   E   E   E   I         224
ATT AGT GGT GAA GAG GAA GAG GAG GAG GAA GAA GAG ATG TTA ATC AGT GAA GAG GAG ATA   672

 P   F   K   D   D   P   R   D   E   T   Y   K   P   H   L   E   R   E   T   P     244
CCA TTC AAA GAT GAT CCA AGA GAT GAG ACC TAC AAA CCC CAC TTA GAA AGG GAA ACC CCA   732

 K   P   R   R   K   S   G   K   V   K   E   E   K   E   K   K   E   I   K   V     264
AAG CCA CGG AGA AAA TCA GGG AAG GTA AAA GAA GAG AAG GAG AAG AAG GAA ATT AAA GTG   792

 E   V   E   V   E   V   K   E   E   E   N   E   I   R   E   D   E   E   P   P     284
GAA GTA GAG GTG GAG GTG AAA GAA GAG GAG AAT GAA ATT AGA GAG GAT GAG GAA CCT CCA   852

 R   K   R   G   R   R   R   K   D   D   K   S   P   R   L   P   K   R   R   K     304
AGG AAG AGA GGA AGA AGA CGA AAA GAT GAC AAA AGT CCA CGT TTA CCC AAA AGG AGA AAA   912

 K   P   P   I   Q   Y   V   R   C   E   M   E   G   C   G   T   V   L   A   H     324
AAG CCT CCA ATC CAG TAT GTC CGT TGT GAG ATG GAA GGA TGT GGA ACT GTC CTT GCC CAT   972

 P   R   Y   L   Q   H   H   I   K   Y   Q   H   L   L   K   K   K   Y   V   C     344
CCT CGC TAT TTG CAG CAC CAC ATT AAA TAC CAG CAT TTG CTG AAG AAG AAA TAT GTA TGT  1032

 P   H   P   S   C   G   R   L   F   R   L   Q   K   Q   L   L   R   H   A   K     364
CCC CAT CCC TCC TGT GGA CGA CTC TTC AGG CTT CAG AAG CAA CTT CTG CGA CAT GCC AAA  1092

 H   H   T   D   Q   R   D   Y   I   C   E   Y   C   A   R   A   F   K   S   S     384
CAT CAT ACA GAT CAA AGG GAT TAT ATC TGT GAA TAT TGT GCT CGG GCC TTC AAG AGT TCC  1152
```

FIG. 1A

```
 H   N   L   A   V   H   R   M   I   H   T   G   E   K   P   L   Q   C   E   I    404
CAC AAT CTG GCA GTG CAC CGG ATG ATT CAC ACT GGC GAG AAG CCA TTA CAA TGT GAG ATC   1212

 C   G   F   T   C   R   Q   K   A   S   L   N   W   H   M   K   K   H   D   A     424
TGT GGA TTT ACT TGT CGA CAA AAG GCA TCT CTT AAT TGG CAC ATG AAG AAA CAT GAT GCA   1272

 D   S   F   Y   Q   F   S   C   N   I   C   G   K   K   F   E   K   K   D   S     444
GAC TCC TTC TAC CAG TTT TCT TGC AAT ATC TGT GGC AAA AAA TTT GAG AAG AAG GAC AGC   1332

 V   V   A   H   K   A   K   S   H   P   E   V   L   I   A   E   A   L   A   A     464
GTA GTG GCA CAC AAG GCA AAA AGC CAC CCT GAG GTG CTG ATT GCA GAA GCT CTG GCT GCC   1392

 N   A   G   A   L   I   T   S   T   D   I   L   G   T   N   P   E   S   L   T     484
AAT GCA GGC GCC CTC ATC ACC AGC ACA GAT ATC TTG GGC ACT AAC CCA GAG TCC CTG ACG   1452

 Q   P   S   D   G   Q   G   L   P   L   L   P   E   P   L   G   N   S   T   S     504
CAG CCT TCA GAT GGT CAG GGT CTT CCT CTT CTT CCT GAG CCC TTG GGA AAC TCA ACC TCT   1512

 G   E   C   L   L   L   E   A   E   G   M   S   K   S   Y   C   S   G   T   E     524
GGA GAG TGC CTA CTG TTA GAA GCT GAA GGG ATG TCA AAG TCA TAC TGC AGT GGG ACG GAA   1572

 R   V   S   L   M   A   D   G   K   I   F   V   G   S   G   S   S   G   G   T     544
CGG GTG AGC CTG ATG GCT GAT GGG AAG ATC TTT GTG GGA AGC GGC AGC AGT GGA GGC ACT   1632

 E   G   L   V   M   N   S   D   I   L   G   A   T   T   E   V   L   I   E   D     564
GAA GGG CTG GTT ATG AAC TCA GAT ATA CTC GGT GCT ACC ACA GAG GTT CTG ATT GAA GAT   1692

 S   D   S   A   G   P   *                                                        570
TCA GAC TCT GCC GGA CCT TAG TGGACAGGAAGACTTGGGGCATGGGACAGCTCAGACTTTGTATTTAAAAGT   1761

TAAAAAGGACAAAAAAAAAAAAAAAAAAAA                                                    1791
```

FIG. 1B

```
                    GCACGAGCGGAGAGCCGCGCAGGGCGCGGGCCGCGCGGGGTGGGGCAGCCGGAGCGCA

            GGCCCCCGATCCCCGGCGGGCGCCCCCGGGCCCCCGCGCGCGCCCCGGCCTCCGGGAGACTGGCGCATGCCACGGAGCG

            CCCCTCGGGCCGCCGCCGCTCCTGCCCGGGCCCCTGCTGCTGCTGCTGTCGCCTGCGCCTGCTGCCCCAACTCGGCGCC

                                                                            M   F          2
            CGACTCACAAAGAAACATCATGTTCGCTCCTTAGCAGGCAAACGACTTTTCTCCTCGCCTCCTCGCCCCGC ATG TTC     6

             R   T   K   R   S   A   L   V   R   R   L   W   R   S   R   A   P   G   G   E    22
            AGG ACC AAA CGA TCT GCG CTC GTC CGG CGT CTC TGG AGG AGC CGT GCG CCC GGC GGC GAG    66

             D   E   E   E   G   A   G   G   G   G   G   G   E   L   R   G   E   G   A         42
            GAC GAG GAG GAG GGC GCA GGG GGA GGT GGA GGA GGA GGC GAG CTG CGG GGA GAA GGG GCG   126

             T   D   S   R   A   H   G   A   G   G   G   G   P   G   R   A   G   C   C   L     62
            ACG GAC AGC CGA GCG CAT GGG GCC GGT GGC GGC GGC CCG GGC AGG GCT GGA TGC TGC CTG   186

             G   K   A   V   R   G   A   K   G   H   H   H   P   H   P   P   A   A   G   A     82
            GGC AAG GCG GTG CGA GGT GCC AAA GGT CAC CAC CAT CCC CAC CCG CCA GCC GCG GGC GCC   246

             G   A   A   G   G   A   E   A   D   L   K   A   L   T   H   S   V   L   K   K    102
            GGC GCG GCC GGG GGC GCC GAG GCG GAT CTG AAG GCG CTC ACG CAC TCG GTG CTC AAG AAA   306

             L   K   E   R   Q   L   E   L   L   L   Q   A   V   E   S   R   G   G   T   R    122
            CTG AAG GAG CGG CAG CTG GAG CTG CTG CTC CAG GCC GTG GAG TCC CGC GGC GGG ACG CGC   366

             T   A   C   L   L   L   P   G   R   L   D   C   R   L   G   P   G   A   P   A    142
            ACC GCG TGC CTC CTG CTG CCC GGC CGC CTG GAC TGC AGG CTG GGC CCG GGG GCG CCC GCC   426

             G   A   Q   P   A   Q   P   P   S   S   Y   S   L   P   L   L   L   C   K   V    162
            GGC GCG CAG CCT GCG CAG CCG CCC TCG TCC TAC TCG CTC CCC CTC CTG CTG TGC AAA GTG   486

             F   R   W   P   D   L   R   H   S   S   E   V   K   R   L   C   C   C   E   S    182
            TTC AGG TGG CCG GAT CTC AGG CAT TCC TCG GAA GTC AAG AGG CTG TGT TGC TGT GAA TCT   546

             Y   G   K   I   N   P   E   L   V   C   C   N   P   H   H   L   S   R   L   C    202
            TAC GGG AAG ATC AAC CCC GAG CTG GTG TGC TGC AAC CCC CAT CAC CTT AGC CGA CTC TGC   606

             E   L   E   S   P   P   P   P   Y   S   R   Y   P   M   D   F   L   K   P   T    222
            GAA CTA GAG TCT CCC CCC CCT CCT TAC TCC AGA TAC CCG ATG GAT TTT CTC AAA CCA ACT   666

             A   D   C   P   D   A   V   P   S   S   A   E   T   G   G   T   N   Y   L   A    242
            GCA GAC TGT CCA GAT GCT GTG CCT TCC TCC GCT GAA ACA GGG GGA ACG AAT TAT CTG GCC   726

             P   G   G   L   S   D   S   Q   L   L   L   E   P   G   D   R   S   H   W   C    262
            CCT GGG GGG CTT TCA GAT TCC CAA CTT CTT CTG GAG CCT GGG GAT CGG TCA CAC TGG TGC   786

             V   V   A   Y   W   E   E   K   T   R   V   G   R   L   Y   C   V   Q   E   P    282
            GTG GTG GCA TAC TGG GAG GAG AAG ACG AGA GTG GGG AGG CTC TAC TGT GTC CAG GAG CCC   846

             S   L   D   I   F   Y   D   L   P   Q   G   N   G   F   C   L   G   Q   L   N    302
            TCT CTG GAT ATC TTC TAT GAT CTA CCT CAG GGG AAT GGC TTT TGC CTC GGA CAG CTC AAT   906

             S   D   N   K   S   Q   L   V   Q   K   V   R   S   K   I   G   C   G   I   Q    322
            TCG GAC AAC AAG AGT CAG CTG GTG CAG AAG GTG CGG AGC AAA ATC GGC TGC GGC ATC CAG   966

             L   T   R   E   V   D   G   V   W   V   Y   N   R   S   S   Y   P   I   F   I    342
            CTG ACG CGG GAG GTG GAT GGT GTG TGG GTG TAC AAC CGC AGC AGT TAC CCC ATC TTC ATC  1026

             K   S   A   T   L   D   N   P   D   S   R   T   L   L   V   H   K   V   F   P    362
            AAG TCC GCC ACA CTG GAC AAC CCG GAC TCC AGG ACG CTG TTG GTA CAC AAG GTG TTC CCC  1086
```

FIG. 2A

```
  G   F   S   I   K   A   F   D   Y   E   K   A   Y   S   L   Q   R   P   N   D        382
GGT TTC TCC ATC AAG GCT TTC GAC TAC GAG AAG GCG TAC AGC CTG CAG CGG CCC AAT GAC       1146

  H   E   F   M   Q   Q   P   W   T   G   F   T   V   Q   I   S   F   V   K   G        402
CAC GAG TTT ATG CAG CAG CCG TGG ACG GGC TTT ACC GTG CAG ATC AGC TTT GTG AAG GGC       1206

  W   G   Q   C   Y   T   R   Q   F   I   S   S   C   P   C   W   L   E   V   I        422
TGG GGT CAG TGC TAC ACC CGC CAG TTC ATC AGC AGC TGC CCG TGC TGG CTA GAG GTC ATC       1266

  F   N   S   R   *                                                                    426
TTC AAC AGC CGG TAG    CCGCGTGCGGAGGGGACAGAGCGTGAGCTGAGCAGGCCACACTTCAAACTACTTTGCT      1278
```

GCTAATATTTTCCTCCTGAGTGCTTGCTTTTCATGCAAACTCTTTGGTCGTTTTTTTTTTGTTTGTTGGTTGGTTTTCT

TCTTCTCGTCCTCGTTTGTGTTCTGTTTTGTTTCGCTCTTTGAGAAATAGCTTATGAAAAGAATTGTTGGGGGTTTTTT

TGGAAGAAGGGGCAGGTATGATCGGCAGGACACCCTGATAGGAAGAGGGGAAGCAGAAATCCAAGCACCACCAAACACA

GTGTATGAAGGGGGGCGGTCATCATTTCACTTGTCAGGAGTGTGTGTGAGTGTGAGTGTGCGGCTGTGTGTGCACGCGT

GTGCAGGAGCGGCAGATGGGGAGACAACGTGCTCTTTGTTTTGTGTCTCTTATGGATGTCCCCAGCAGAGAGGTTTGCA

GTCCCAAGCGGTGTCTCTCCTGCCCCTTGGACACGCTCAGTGGGGCAGAGGCAGTACCTGGGCAAGCTGGCGGCTGGGG

TCCCAGCAGCTGCCAGGAGCACGGCTCTGTCCCCAGCCTGGGAAAGCCCCTGCCCCTCCTCTCCCTCATCAAGGACACG

GGCCTGTCCACAGGCTTCTGAGCAGCGAGCCTGCTAGTGGCCGAACCAGAACCAATTATTTTCATCCTTGTCTTATTCC

CTTCCTGCCAGCCCCTGCCATTGTAGCGTCTTTCTTTTTTGGCCATCTGCTCCTGGATCTCCCTGAGATGGGCTTCCCA

AGGGCTGCCGGGGCAGCCCCCTCACAGTATTGCTCACCCAGTGCCCTCTCCCCTCAGCCTCTCCCCTGCCTGCCCTGGT

GACATCAGGTTTTTTCCCGGACTTAGAAAACCAGCTCAGCACTGCCTGCTCCCATCCTGTGTGTTAAGCTCTGCTATTAG

GCCAGCAAGCGGGGATGTCCCTGGGAGGGACATGCTTAGCAGTCCCCTTCCCTCCAAGAAGGATTTGGTCCGTCATAAC

CCAAGGTACCATCCTAGGCTGACACCTAACTCTTCTTTCATTTCTTCTACAACTCATACACTCGTATGATACTTCGACA

CTGTTCTTAGCTCAATGAGCATGTTTAGACTTTAACATAAGCTATTTTTCTAACTACAAAGGTTTAAATGAACAAGAGA

AGCATTCTCATTGGAAATTTAGCATTGTAGTGCTTTGAGAGAGAAAGGACTCCTGAAAAAAAACCTGAGATTTATTAAA

GAAAAAAATGTATTTTATGTTATATATAAATATATTATTACTTGTAAATATAAAGACGTTTTATAAGCATCATTATTTA

TGTATTGTGCAATGTGTATAAACAAGAAAAATAAAGAAAAGATGCACTTTGCTTTAATATAAATGCAAATAACAAATGC

CAAATTAAAAAAGATAAACACAAGATTGGTGTTTTTTCCTATGGGTGTTATCACCTAGCTGAATGTTTTTCTAAAGGAG

TTTATGTTCCATTAAACGATTTTTAAAAATGTACACTTGAAAAAAAAAAAAAAAAAA

<div align="center">FIG. 2B</div>

GGCACGAGGTTGCCCTGGCGGAGCAGAGACAGGCCCTCGGGGTGGAGGTC

```
                                        M   C   N   T   P   T   Y   C   D   L    10
TTTGGTTTCATAAGAGCCTGAGAGAGATTTTTCTAAGAT ATG TGT AAC ACA CCA ACG TAC TGT GAC CTA   30

  G   K   A   A   K   D   V   F   N   K   G   Y   G   F   G   M   V   K   I   D    30
GGA AAG GCT GCT AAG GAT GTC TTC AAC AAA GGA TAT GGC TTT GGC ATG GTC AAG ATA GAC   90

  L   K   T   K   S   C   S   G   V   E   F   S   T   S   G   H   A   Y   T   D    50
CTG AAA ACC AAG TCT TGT AGT GGA GTG GAA TTT TCT ACT TCT GGT CAT GCT TAC ACT GAT   150

  T   G   K   A   S   G   N   L   E   T   K   Y   K   V   C   N   Y   G   L   T    70
ACA GGG AAA GCA TCA GGC AAC CTA GAA ACC AAA TAT AAG GTC TGT AAC TAT GGA CTT ACC   210

  F   T   Q   K   W   N   T   D   N   T   L   G   T   E   I   S   W   E   N   K    90
TTC ACC CAG AAA TGG AAC ACA GAC AAT ACT CTA GGG ACA GAA ATC TCT TGG GAG AAT AAG   270

  L   A   E   G   L   K   L   T   L   D   T   I   F   V   P   N   T   G   K   K   110
TTG GCT GAA GGG TTG AAA CTG ACT CTT GAT ACC ATA TTT GTA CCG AAC ACA GGA AAG AAG   330

  S   G   K   L   K   A   S   Y   K   R   D   C   F   S   V   G   S   N   V   D   130
AGT GGG AAA TTG AAG GCC TCC TAT AAA CGG GAT TGT TTT AGT GTT GGC AGT AAT GTT GAT   390

  I   D   F   S   G   P   T   I   Y   G   W   A   V   L   A   F   E   G   W   L   150
ATA GAT TTT TCT GGA CCA ACC ATC TAT GGC TGG GCT GTG TTG GCC TTC GAA GGG TGG CTT   450

  A   G   Y   Q   M   S   F   D   T   A   K   S   K   L   S   Q   N   N   F   A   170
GCT GGC TAT CAG ATG AGT TTT GAC ACA GCC AAA TCC AAA CTG TCA CAG AAT AAT TTC GCC   510

  L   G   Y   K   A   A   D   F   Q   L   H   T   H   V   N   D   G   T   E   F   190
CTG GGT TAC AAG GCT GCG GAC TTC CAG CTG CAC ACA CAT GTG AAC GAT GGC ACT GAA TTT   570

  G   G   S   I   Y   Q   K   V   N   E   K   I   E   T   S   I   N   L   A   W   210
GGA GGT TCT ATC TAC CAG AAG GTG AAT GAG AAG ATT GAA ACA TCC ATA AAC CTT GCT TGG   630

  T   A   G   S   N   N   T   R   F   G   I   A   A   K   Y   M   L   D   C   R   230
ACA GCT GGG AGT AAC AAC ACC CGT TTT GGC ATT GCT GCT AAG TAC ATG CTG GAT TGT AGA   690

  T   S   L   S   A   K   V   N   N   A   S   L   I   G   L   G   Y   T   Q   T   250
ACT TCT CTC TCT GCT AAA GTA AAT AAT GCC AGC CTG ATT GGA CTG GGT TAT ACT CAG ACC   750

  L   R   P   G   V   K   L   T   L   S   A   L   I   D   G   K   N   F   S   A   270
CTT CGA CCA GGA GTC AAA TTG ACT TTA TCA GCT TTA ATC GAT GGG AAG AAC TTC AGT GCA   810

  G   G   H   K   V   G   L   G   F   E   L   E   A   *                           283
GGA GGT CAC AAG GTT GGC TTG GGA TTT GAA CTG GAA GCT TAA TGTGGTTTGAGGAAAGCATCAGA   849
```

TTTGTCCCTGGAAGTGAAGAGAAATGAACCCACTATGTTTTGGCCTTAAAATTCTTCTGTGAAATTTCAAAAGTGTGAA

CTTTTTATTCTTCCAAAGAATTGTAATCCTCCCCACACTGAAGTCTAGGGGTTGCGAATCCCTCCTGAGGGAGACGCTT

GAAGGCATGCCTGGAAGTTGTCATGTTTGTGCCACGTTTCAGTTCAGTTCTGAAGTGTTATTAAATGTGTTCCTCAGCG

ACAGTGTAGCGTCATGTTAGAGGAGACGATCTGACCCACCAGTTTGTACATCACGTCCTGCATGTCCCACACCATTTTT

TCATGACCTTGTAATATACTGGTCTCTGTGCTATAGTGGAATCTTTGGTTTTTGCATCATAGTAAAATAAAATAAACCCA
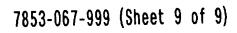
TCACATTTGGAACATAAAAAAAAAAAAAAAAAAAAA

FIG. 3

```
 T   S   L   A   L   V   L   N   L   L   Q   I   Q   R   N   V   T   L   F   P
ACG AGC CTA GCC CTG GTG CTC AAC CTG CTG CAG ATC CAG AGG AAT GTC ACT CTC TTC CCC      20
                                                                                     60

 E   E   V   I   A   T   I   F   S   S   A   W   W   V   P   P   C   C   G   T
GAG GAG GTG ATC GCC ACC ATC TTT TCC TCC GCC TGG TGG GTC CCT CCC TGC TGC GGG ACA      40
                                                                                    120

 A   A   A   V   V   G   L   L   Y   P   C   I   D   S   H   L   G   E   P   H
GCA GCT GCT GTT GTT GGC CTA CTG TAC CCC TGT ATC GAC AGT CAC CTC GGA GAA CCC CAC      60
                                                                                    180

 K   F   K   R   E   W   A   S   V   M   R   C   I   A   V   F   V   G   I   N
AAA TTT AAG AGA GAA TGG GCC AGT GTC ATG CGC TGC ATA GCA GTT TTT GTT GGC ATT AAC      80
                                                                                    240

 H   A   S   A   K   L   D   F   A   N   N   V   Q   L   S   L   T   L   A   A
CAC GCC AGT GCT AAA TTG GAT TTT GCC AAT AAT GTC CAG CTG TCC TTG ACT TTA GCA GCC     100
                                                                                    300

 L   S   L   G   L   W   W   T   F   D   R   S   R   S   G   L   G   L   G   I
CTA TCT TTG GGC CTT TGG TGG ACA TTT GAT CGT TCC AGA AGT GGC CTT GGG CTG GGG ATC     120
                                                                                    360

 T   I   A   F   L   A   T   L   I   T   Q   F   L   V   Y   N   G   V   Y   Q
ACC ATA GCT TTT CTA GCT ACG CTG ATC ACG CAG TTT CTC GTG TAT AAT GGT GTC TAT CAG     140
                                                                                    420

 Y   T   S   P   D   F   L   Y   I   R   S   W   L   P   C   I   F   F   S   G
TAT ACA TCC CCA GAT TTC CTC TAT ATT CGT TCT TGG CTC CCT TGT ATA TTT TTC TCA GGA     160
                                                                                    480

 G   V   T   V   G   N   I   G   R   Q   L   A   M   G   V   P   E   K   P   H
GGC GTC ACG GTG GGG AAC ATA GGA CGA CAG TTA GCT ATG GGT GTT CCT GAA AAG CCC CAT     180
                                                                                    540

 S   D   *
AGT GAT TGA GTCTTCAAAACCACCGATTCTGAGAGCAAGGAAGATTTTGGAAGAAAATCTGACTGTGGATTATGAC     182
                                                                                    546

AAAGATTATCTTTTTTCTTAAGTAATCTATTTAGATCGGGCTGACTGTACAAATGACTCCTGGAAAAAACTCTTCACCT

AGTCTAGAATAGGGAGGTGGAGAATGATGACTTACCCTGAAGTCTTCCCTTGACTGCCCGCACTGGCGCCTGTCTGTGC

CCTGGAGCATTCTGCCCAGGCTACGTGGGTTCAGGCAGGTGGCAGCTTCCCAAGTATTCGATTTCATTCATGTGATTAA

AACAAGTTGCCATATTTCAAAAAAAAAAAAAAAAAAAMCTCGAGACCAACCCGCAGTTTTGTGTCAGTGCCCAAAGGAGGT

AGGTTGATGGTGCTTAACAAACATGAAGTATGGTGTAATAGGAATAATATTTATCCNAAAGATTTTTAAAAAATAGGGCT

GTGTTTAAAAAAAAAAAAAAAAAAAAAA
```

FIG. 4

```
M   C   H   S   R   S   C   H   P   T   M   T   I   L   Q   A   P   T   P   A    20
ATG TGT CAC TCT CGC AGC TGC CAC CCG ACC ATG ACC ATC CTG CAG GCC CCG ACC CCG GCC  60

P   S   T   I   P   G   P   R   R   G   S   G   P   E   I   F   T   F   D   P    40
CCC TCC ACC ATC CCG GGA CCC CGG CGG GGC TCC GGT CCT GAG ATC TTC ACC TTC GAC CCT  120

L   P   E   P   A   A   A   P   A   G   R   P   S   A   S   R   G   H   R   K    60
CTC CCG GAG CCC GCA GCG GCC CCT GCC GGG CGC CCC AGC GCC TCT CGC GGG CAC CGA AAG  180

R   S   R   R   V   L   Y   P   R   V   V   R   R   Q   L   P   V   E   E   P    80
CGC AGC CGC AGG GTT CTC TAC CCT CGA GTG GTC CGG CGC CAG CTG CCA GTC GAG GAA CCG  240

N   P   A   K   R   L   L   F   L   L   L   T   I   V   F   C   Q   I   L   M    100
AAC CCA GCC AAA AGG CTT CTC TTT CTG CTG CTC ACC ATC GTC TTC TGC CAG ATC CTG ATG  300

A   E   E   G   V   P   A   P   L   P   P   E   D   A   P   N   A   A   S   L    120
GCT GAA GAG GGT GTG CCG GCG CCC CTG CCT CCA GAG GAC GCC CCT AAC GCC GCA TCC CTG  360

A   P   T   P   V   S   P   V   L   E   P   F   N   L   T   S   E   P   S   D    140
GCG CCC ACC CCT GTG TCC CCC GTC CTC GAG CCC TTT AAT CTG ACT TCG GAG CCC TCG GAC  420

Y   A   L   D   L   S   T   F   L   Q   Q   H   P   A   A   F   *              157
TAC GCT CTG GAC CTC AGC ACT TTC CTC CAG CAA CAC CCG GCC GCC TTC TAA             471

CTGTGACTCCCCGCACTCCCCAAAAAGAATCCGAAAAACCACAAAGAAACACCAGGCGTACCTGGTGCGCGAGAGCGTA  550

TCCCCAACTGGGACTTCCGAGGCAACTTGAACTCAGAACACTACAGCGGAGACGCCACCCGGTGCTTGAGGCGGGACCG  629

AGGCGCACAGAGACCGAGGCGCATAGAGACCGAGGCACAGCCCAGCTGGGGCTAGGCCCGGTGGGAAGGAGAGCGTCGT  708

TAATTTATTTCTTATTGCTCCTAATTAATATTTATATGTATTTATGTACGTCCTCCTAGGTGATGGAGATGTGTACGTA  787

ATATTTATTTTAACTTATGCAAGGGTGTGAGATGTTCCCTCTGCTGTAAATGCAGGTCTCTTGGTATTTATTGAGCTTT  866

GTGGGACTGGTGGAAGCAGGACACCTGGAACTGCGGCAAAGTAGGAGAAGAAATGGGGAGGACTCGGGTGGGGGAGGAC  945

GTCCCGGCTGGGATGAAGTCTGGTGGTGGGTCGTAAGTTTAGGAGGTGACTGCATCCTCCAGCATCTCAACTCCGTCTG  1024

TCTACTGTGTGAGACTTCGGCGGACCATTAGGAATGAGATCCGTGAGATCCTTCCATCTTCTTGAAGTCGCCTTTAGGG  1103

TGGCTGCGAGGTAGAGGGTTGGGGGTTGGTGGGCTGTCACGGAGCGACTGTCGAGATCGCCTAGTATGTTCTGTGAACA  1182

CAAATAAAATTGATTTACTGTCAAAAAAAAAAAAAAAAAAAACTCGAG                                1228
```

FIG. 5

GAATTCGGCACGAGGHCAGGAGCTCCTTTTACTGCGTCTCCCATCATGGGCTTAGGGTTGAGTCTTCA 68

GGTTCTGGGGGCAGGAAGGACGGGCACTCAGGAGCCCCCCTCCCCATCCACAGCCCCTCTTTGGGAGGGGGGAAACTTG 147

GCAACCCGGGAGGCATGTGGATCTTTTCCTAAGCAAGATGCTGAGCTCGAAGATGGGGGTGTAAGGTAATGTCCCAAA 226

CTGAAACTTTGCCAGGCACTGGGAGAGGCTGTGAACTCTTTTCTGGCTTTAGAATTTAGGTCTAGATCCCAAAAGGCTAT 305

AGTACCCCCTGGGGGCTAACCAGAGGCATGCCTGGGCTGAGCTGAACCTTCTGGTGCACTGGCCCCTGGCTGACTGCTC 384

TTCTGCAGGAAGTTGGAGGAGATTCCTGAAGTTGATTCCTCAGGCTGGATGTCCAAGGGGGTTGGAGTTTCTGATGTCT 463

TTCTGTCTCCCTCTCTTTTCTTTCTCTCCCTACCAGGTCCACTTCTTTCAGAGGGGCCTGCGGTGCTCTAAAAGTTCTC 542

CTGTTAAAGTTTAGAGCAAATTGGTTATTATTTTAAAATCAATAAAACTTTTAAAAGTACTAAGACAACTTCTAAGAGG 621

GGAGTGGACAGAGGGCCTGGTGGCAGCTCACAGTTTCTTTTCTGACCTTTGGTCTCACCCACCAAGTGTCCCACCTGAG 700

TGCCCACCTTGCCCACCTGAGGTAATGCCCTGGGGCTCCACCAGTCCAGATCCACAGGGCGCAGCCATGTGGGAGTGGC 779

GGCTGATTGTTACCCAGTAGTGTTGATAGCACATTATTCATAACAGCCAAAGAGAGGAAGCAACCCAAATGTCCATTAG 858

CTGATAAATGGATAAATGAAATATGGTACGTCCGAAGAATGGAATATCATTCACCCATGAAAAAGAACGAAGTCCAGCA 937

CCAAAACGTGCTACAACATGGATGAACTTCGATGACTTTGTGCCACATGAAAGAAGAAGCCAGCCACAAAAGGCCATAT 1016

```
         M   S   R   M   G   K   P   I   E   T   Q   K   S   P   P   P    16
ATTGTATGAAATGAA ATG TCC AGA ATG GGC AAA CCC ATA GAG ACA CAA AAA TCT CCG CCA CCT 1079

 P   Y   S   R   L   S   P   R   D   E   Y   K   P   L   D   L   S   D   S . T    36
CCC TAC TCT CGG CTG TCT CCT CGC GAC GAG TAC AAG CCA CTG GAT CTG TCC GAT TCC ACA 1139

 L   S   Y   T   E   T   E   A   T   N   S   L   I   T   A   P   G   E   F   S    56
TTG TCT TAC ACT GAA ACG GAG GCT ACC AAC TCC CTC ATC ACT GCT CCG GGT GAA TTC TCA 1199

 D   A   S   M   S   P   D   A   T   K   P   S   H   W   C   S   V   A   Y   W    76
GAC GCC AGC ATG TCT CCG GAC GCC ACC AAG CCG AGC CAC TGG TGC AGC GTG GCG TAC TGG 1259

 E . H   R   T   R   V   G   R   L   Y   A   V   Y   D   Q   A   V   S   I   F    96
GAG CAC CGG ACG CGC GTG GGC CGC CTC TAT GCG GTG TAC GAC CAG GCC GTC AGC ATC TTC 1319

 Y   D   L   P   Q   G   S   G   F   C   L   G   Q   L   N   L   E   Q   R   S   116
TAC GAC CTA CCT CAG GGC AGC GGC TTC TGC CTG CGC CAG CTC AAC CTG GAG CAG CGC AGC 1379

 E   S   V   R   R   T   R   S   K   I   G   F   G   I   L   L   S   K   E   P   136
GAG TCG GTG CGG CGA ACG CGC AGC AAG ATC GGC TTC GGC ATC CTG CTC AGC AAG GAG CCC 1439

 D   G   V   W   A   Y   N   R   G   E   H   P   I   F   V   N   S   P   T   L   156
GAC GGC GTG TGG GCC TAC AAC CGC GGC GAG CAC CCC ATC TTC GTC AAC TCC CCG ACG CTG 1499

 D   A   P   G   G   R   A   L   V   V   R   K   V   P   P   G   Y   S   I   K   176
GAC GCG CCC GGC GGC CGC GCC CTG GTC GTG CGC AAG GTG CCC CCC GGC TAC TCC ATC AAG 1559

 V   F   D   F   E   R   S   G   L   Q   H   A   P   E   P   D   A   A   D   G   196
GTG TTC GAC TTC GAG CGC TCG GGC CTG CAG CAC GCG CCC GAG CCC GAC GCC GCC GAC GGC 1619

 P   Y   D   P   N   S   V   R   I   S   F   A   K   G   W   G   P   C   Y   S   216
CCC TAC GAC CCC AAC AGC GTC CGC ATC AGC TTC GCC AAG GGC TGG GGG CCC TGC TAC TCC 1679

 R   Q   F   I   T   S   C   P   C   W   L   E   I   L   L   N   N   P   R   •   235
CGG CAG TTC ATC ACC TCC TGC CCC TGC TGG CTG GAG ATC CTC CTC AAC AAC CCC AGA TAG 1739
```

FIG. 6A

```
TGGCGGCCCCGGCGGGAGGGGCCGGGTGGGACGCCGCGGCCACCGCCACCTGCCGGCCTCGAGAGGGGCCGATGCCCAGA   1818

GACACAGCCCCCACGGACAAAACCCCCCAGATATCATCTACCTAGATTTAATATAAAGTTTTATATATTATATGGAAAT   1897

ATATATTATACTTGTAATTATGGAGTCATTTTTACAATGTAATTATTTATGTATGGTGCAATGTGTGTATATGGACAAA   1976

ACAAGAAAGACGCACTTTGGCTTATAATTCTTTCAATACAGATATATTTTCTTTCTCTTCCTCCTTCCTCTTCCTTACT   2055

TTTTATATATATATATAAAGAAAATGATACAGCAGAGCTAGGTGGAAAAGCCTGGGTTTGGTGTATGGTTTTTGAGATA   2134

TTAATGCCCAGACAAAAAGCTAATACCAGTCACTCGATAATAAAGTATTCGCATTATAGTTTTTTTTTAAACTGTCTTCT   2213

TTTTACAAAGAGGGGCAGGTAGGGCTTCAGCGGATTTCTGACCCATCATGTACCTTGAAACTTGACCTCAGTTTTCAAG   2292

TTTTACTTTTATTGGATAAAGACAGAACAAATTGAAAAGGGAGGAAAGTCACATTTACTCTTAAGTAAACCAGAGAAAG   2371

TTCTGTTGTTCCTTCCTGCCCATGGCTATGGGGTGTCCAGTGGATAGGGATGGCGGTGGGGAAAAGGAGAATACACTGG   2450

CCATTTATCCTGGACAAGCTCTTCCAGTCTGATGGAGGAGGTTCATGCCCTAGCCTAGAAAGGCCCAGGTCCATGACCC   2529

CCATCTTTGAGTTATGAGCAAGCTAAAAGAAGACACTATTTCTCACCATTTTGTGGAAATGGCCTGGGGAACAAAGACT   2608

GAAATGGGCCTTGAGCCCACCTGCTACCTTGCAGAGAACCATCTCGAGCCCCGTAGATCTTTTTAGGACCTCCACAGGC   2687

TATTTCCCACCCCCCAGCCAAAAATAGCTCAGAATCTGCCCATCCAGGGCTGTATTAATGATTTATGTAAAGGCAGATG   2766

GTTTATTTCTACTTTGTAAAAGGGAAAAGTTGAGGTTCTGGAAGGATAAATGATTTGCTCATGAGACAAAATCAAGGTT   2845

AGAAGTTACATGGAATTGTAGGACCAGAGCCATATCATTAGATCAGCTTTCTGAAGAATATTCTCHAAAAAAGAAAGTC   2924

TCCTTGGCCAGATAACTAAGAGGAATGTTTCATTGTATATCTTTTTTCTTGGAGATTTATATTAACATATTAAGTGCTC   3003

TGAGAAGTCCTGTGTATTATCTCTTGCTGCATAATAAATTATCCCCAAACTTAAAAAAAAAAAAAAAAAAAAAAAACTCGA   3082

G                                                                                 3083
```

FIG. 6B